Ein Verlagsbeihefter der Heise Zeitschriften Verlag GmbH & Co. KG

iX extra Networking

Schwerpunkt: Load Balancing

Veranstaltungen

10. – 11. September, Düsseldorf
IP Multimedia Subsystem
www.euroforum.de/inno-IMS08

18. September, Düsseldorf
25. September, München
1. Oktober, Hamburg
Heise-Netze-Konferenz 2008:
E-Mail – Sicherheit, Mobilität, Archivierung
www.heise-events.de

21. – 22. Oktober, München
Systems 2008 mit Fachmessen Communication World 2008 und Internet World 2008
www.systems.de

27. – 29. Oktober, Frankfurt am Main
Storage Networking World Europe 2008
www.snweurope.com

iX extra
Networking zum Nachschlagen:
www.heise.de/ix/extra/netzwerke.shtml

Des anderen Last

Load Balancing beschleunigt Server-Anwendungen

Wie schaffen es Unternehmen mit stark frequentierten Websites, auch bei Zehntausenden gleichzeitigen Nutzern die Antwortzeiten kurz zu halten? Vor allem durch zahlreiche Server und ausgefeilte Verfahren, die die Last der Anfragen sinnvoll verteilen.

Trotz nach wie vor zunehmender Performance von CPUs und I/O-Subsystemen gibt es immer wieder Aufgaben, die einen einzelnen Server schlicht überfordern. Dann helfen auch keine Mehrprozessorserver mit schnellem Storage und speziell optimierten Betriebssystemen mehr. Zwar bieten Hersteller inzwischen Server mit zwei- oder dreistelliger CPU-Zahl von der Stange an, doch sind die für bestimmte Rechenaufgaben – etwa fürs High Performance Computing (HPC) – konzipiert und nicht für so banale Aufgaben wie Web- oder Maildienste.

In solchen Fällen bleibt einem nichts anderes übrig, als die Aufgaben auf mehrere Server zu verteilen – was unter Umständen auch preisgünstiger sein kann. Dabei unterscheidet man zwei Formen von Server-Ansammlungen: Während man unter einer Serverfarm einen lose zusammengewürfelten Haufen unabhängiger Server versteht, die nicht unbedingt zusammenarbeiten müssen, ist ein Cluster (Schwarm) wesentlich strikter organisiert. Jeder Server – mit eigenem Betriebssystem – darin bildet einen Knoten oder Node, doch entscheidend ist die Software, die den Cluster zusammenhält, meist eine Middleware, die die Aufgaben auf die verfügbaren Ressourcen verteilt.

Cluster beginnen bei zwei aus Verfügbarkeitsgründen zusammengeschalteten Servern, von denen einer schlummernd auf seinen Einsatz wartet (active/passive), oder die parallel laufen (active/active). Beide fallen häufig unter den Begriff HA-Cluster (High Availability). Das andere Ende bilden als Rechen-Cluster gebaute Supercomputer, deren Spitze momentan der frisch entwickelte und hauptsächlich militärisch genutzte Hybrid-Cluster „Roadrunner“ mit 122 400 Prozessorkernen darstellt. Bei ihnen steht weniger die Verfügbarkeit als vielmehr die akkumulierte Rechenleistung im Vordergrund. Zwischen diesen beiden Extremen existieren so ziemlich alle Formen von Clustern, die man sich vorstellen kann.

Während Serverfarmen häufig einen ganzen Zoo an Servern subsumieren, bestehen Cluster meist aus Rechnern mit identischer Hardware, was die Organisation und Verteilung der Anfragen auf die Server verein-

facht; dies ist aber nicht zwingend notwendig. Das Verteilen von Aufgaben auf mehrere Rechner erhöht grundsätzlich die Verfügbarkeit des Gesamtsystems. Fällt einer aus, stehen die anderen weiterhin zur Verfügung. Außerdem lassen sich Server temporär zum Aufrüsten oder für Wartungsarbeiten aus dem Verbund nehmen. Deshalb bemisst man die Kapazität von Clustern häufig so, dass sie auch beim Ausfall eines Knotens die Lastspitzen abfangen können.

Besitzen die Server interne, exklusive Festplatten, müssen sie sicherstellen, dass sie stets mit dem gleichen Datenbestand arbeiten, also die Daten laufend synchronisieren. Dafür sorgt meist eine Überwachungssoftware.

Gemeinschaftsablage

Übersteigt der Aufwand für das Spiegeln den Nutzdatenverkehr, bieten sich gemeinsam genutzte Speicher an. Voraussetzung dafür ist ein externer Speicher, auf den alle Server Zugriff haben. Bei HA-Clustern genügt meist ein externes Dual-Host-SCSI- oder -SAS-System (Serial Attached SCSI) – samt den entsprechenden Mechanismen, die konkurrierende Zugriffe verhindern.

ANBIETER VON LOAD-BALANCING-PRODUKTEN

Die folgende Übersicht erhebt keinen Anspruch auf Vollständigkeit.

Anbieter	URL
3Com	www.3com.de
Akamai	www.akamai.de
Alcatel-Lucent	www.alcatel-lucent.de
Allot Communications	www.allot.com
Array Networks	www.arraynetworks.net
Astrocom	www.astrocorp.com
Barracuda Networks	www.barracudanetworks.com
CAI Networks	www.cainetworks.com
Celestix	www.celestix.com
Cisco Systems	www.cisco.de
Citrix	www.citrix.de
Coyote Point	www.coyotepoint.com
EMC	www.emc.com
F5 Networks	www.f5-networks.com
Foundry Networks	www.foundrynet.com
Hewlett-Packard	www.hewlett-packard.de
Inlab	www.inlab.de
KEMP/MR Systeme	www.load-balancer.info
Loadbalancer.org	www.loadbalancer.org
Microsoft	www.microsoft.de
NENTEC	www.nentec.de
Nortel Networks	www.nortel.de
Oracle	www.oracle.de
Radware	www.radware.com
Resonate	www.resonate.com
Sun Microsystems	www.sun.de
Teamix	www.teamix.de
Telco Tech	www.telco-tech.de
Thomas-Krenn	www.thomas-krenn.de
XGforce.com	www.xgforce.com
Xtera	www.xtera-ip.com
Zeus	www.zeus.com

Die einfachste Form des sogenannten Shared Storage bilden Netzwerkfreigaben per NFS (Unix) oder CIFS (Windows). Darüber können alle autorisierten Systeme auf die benötigten Verzeichnisse und Dateien zugreifen (dateibasiert); für die Datenkonsistenz sorgen NFS und CIFS. Diesem klassischen Fileserver-Prinzip haben die Storage-Hersteller heute meist das – moderner klingende – Etikett NAS (Network Attached Storage) angeklebt.

Andere Cluster und Serverfarmen wiederum sind blockbasiert über ein separates Netz, das Storage Area Network (SAN), mit Speicher-Subsystemen verbunden. Traditionell setzt man bei der Verbindungstechnik auf Fibre Channel (FC), das inzwischen bei 4 respektive 8 GBit/s angelangt ist.

Einige Umgebungen eignen sich aber auch für das sogenannte Internet SCSI oder iSCSI über eigene Gigabit-Ethernet-Leitungen (GE), in seltenen Fällen über 10 GE. Die vielbeschworene Kombination von Ethernet und Fibre Channel hat hier allerdings keinen Platz: Fibre Channel over IP (FCIP) stellt lediglich einen Tunnel zwischen zwei entfernten SANs her, Internet Fibre Channel (iFCP) kann als Hybrid-Protokoll mehr als zwei SANs miteinander verbinden, und Fibre Channel over Ethernet (FCoE) steckt noch im Rohentwurf und benötigt – ebenso wie FCIP und iFCP – spezielle Switches.

Für den Serverzugriff erstellt man auf den im SAN befindlichen Disk-Subsystemen über mehrere Festplatten hinweg Volumes – ähnlich einem lokalen Logical Volume Manager (LVM) – und weist ihnen eine SCSI-LUN (Logical Unit Number) zu. Darauf kann ein autorisierter Server über seinen FC-HBA (Host Bus Adapter), seinen iSCSI-HBA oder seine NIC (Network Interface Card) wie auf eine lokale SCSI-Platte zugreifen. Dazu kommuniziert er mit dem Volume – respektive dem davorsitzenden RAID-Controller – per SCSI-Kommandos.

Das heißt aber auch, dass im Regelfall nur ein Server auf ein Volume zugreifen kann – für Cluster und Serverfarmen benötigt man spezielle Lösungen. Sie bestehen meist aus Cluster-Dateisystemen wie Red Hats Global File System 2 (GFS2), IBMs General Parallel File System (GPFS), Lustre oder OCFS2 (Oracle Cluster File System) [1, 2, 3].

Will man die Speichersysteme ihrerseits zusammenfassen und gemeinsam arbeiten lassen, greift man dagegen auf globale oder verteilte Dateisysteme zurück, die sich über mehrere Speichersubsysteme erstrecken. Darauf greifen mehrere Server oder Cluster-Knoten meist über NFS zu.

Globalisierung

Für bestimmte Anwendungen wie datenbankgestützte Webservices lassen sich mehrstufige Servicekonzepte nutzen, indem man die Systeme aufteilt etwa in Web-, Application- und Datenbankserver. Ersterer nimmt die Anfragen der Clients entgegen und beantwortet sie. Per Skript in die Webseiten eingebaute Datenbankanfragen richtet der beauftragte Application- an den Datenbankserver und reicht die aufbereiteten Daten an den Webserver zurück. Eine große Website kann man außerdem dadurch über mehrere Server verteilen, dass man den Verzeichnisbaum der Site aufsplittet und die Seiten untereinander entsprechend verlinkt.

Ebenso kann man die Last von Webservern mit einer ganz anderen Technik verteilen: dem DNS Load Balancing. Die für eine Zone zuständigen Nameserver lösen in der Regel Host-Namen in Adressen (etwa *www.ix.de* zu *193.99.144.80*) und vice versa auf (siehe auch iX extra 5/2007 [5]), können aber auch einen oder mehrere Redirects von Domain-Namen



Beschleunigte Applikationszustellung.

Endlich: Geschwindigkeit ohne Umschweife.

Durch Ausgleichen des Lastenaufkommens all Ihrer unternehmenskritischen Anwendungen erzielen Sie 100% Verfuegbarkeit und die besten Ansprechzeiten.

Coyote Point hat auf dem Gebiet der preisguenstigen Loesungen zur Applikationsbeschleunigung und -verfuegbarkeit als Marktfuehrer die Nase vorn. Unsere Equalizer TM Produktfamilie bietet hochleistungsfaehigen Lastenausgleich, Applikationsbeschleinigung und Webkompression in einer einzigen, einfach zu verwaltenden Anwendung. Hohe Verfuegbarkeit und hochleistungsfaehige Anwendungszustellung im Handumdrehen.

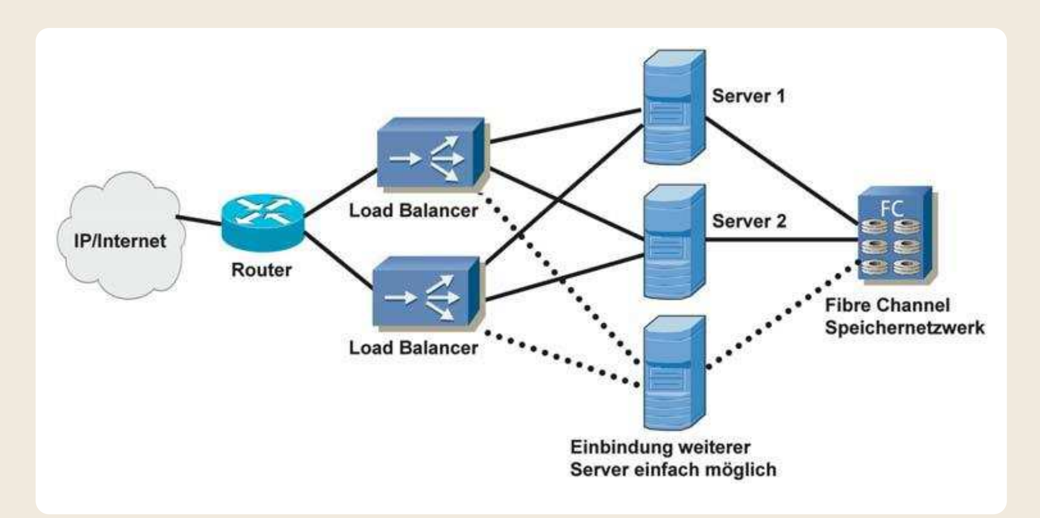
Wenn es um Leistungsfaehigkeit geht, werden Sie keine schnellere, zuverlaessigere oder preiswertere Loesung finden.

Bereits seit ueber zehn Jahren ist Coyote Point in Nordamerika anerkannter Marktfuehrer fuer preisguenstigen, hochleistungsfaehigen Lastenausgleich und Beschleunigung von Applikationen. Wir freuen uns, kuenftig auch Unternehmen der E.U. die gleiche Leistungsfaehigkeit und Qualitaet anbieten zu koennen.

Erfahren Sie mehr darueber, warum ueber 8.000 Unternehmen von den Vorteilen der Coyote Point Loesungen profitieren.

Kontaktieren Sie uns per email unter info@coyotepoint.com oder rufen Sie uns an: 1-877-367-2696.

Im Tierreich ist der Kojote (englisch: Coyote) ein Wesen hoechster Agilitaet und eleganter Einfachheit. Das gleiche sollte fuer Ihr Netzwerk gelten.



Zwei Load Balancer verteilen die Aufgaben auf die vorhandenen Server. Zusätzliche Server lassen sich jederzeit hinzuschalten (Abb. 1).

auf Host-Adressen beinhalten, häufig auf die des *www*-Hosts.

Genauso wie ein Rechner mit einer IP-Adresse auf mehrere Namen (Aliase) hören kann, kann ein Name – egal ob Domain- oder Host-Name – mehreren IP-Adressen und damit mehreren Servern zugeordnet sein. Der Nameserver vergibt die unterschiedlichen IP-Adressen dann entweder reihum oder nach dem Zufallsprinzip und verteilt damit die Zugriffe auf mehrere Server. Durch eine Gewichtung kann er leistungsfähigere oder mit höherem Durchsatz angeschlossene Server stärker belasten.

Dadurch können die Server in den Rechenzentren unterschiedlicher Hoster an mehreren Orten stehen (Global Server Load Balancing, GSLB). Die Methode eignet sich besonders gut zur Erhöhung der Verfügbarkeit. Außerdem liegt ihr Reiz in der einfachen Realisierbarkeit ohne zusätzliche Hardware. Für umfangreiche Installationen und hohe Ansprüche an die Performance eignet sich die Lösung aber weniger – unter anderem, weil der DNS-Server keine Rückmeldungen über die Auslastung der einzelnen Server erhält und somit nur eine statische Lastverteilung stattfindet. Außerdem lassen sich damit nur Internetdienste verteilen. Dafür kann man über eine geschickte Zuordnung der DNS-Einträge den Zugriff auf weltweit verteilte Server optimieren und die Anfragen jeweils dem nächstgelegenen Server zuordnen.

Alternativ arbeiten global operierende Firmen mit weltweit verteilten Caching-Strukturen spezieller Dienstleister wie Akamai. Es liegt in der Natur der Sache, dass bei DNS-gestütztem Load Balancing die Tücken persistenter Sessions oder der SSL-Verschlüsselung (siehe unten) keine Rolle spielen. Da die DNS-Abfrage nur einmal erfolgt und die Antwort für eine gewisse Zeit lokal vorliegt, erreichen alle Folgepakete automatisch denselben Server.

Im Takt reihum

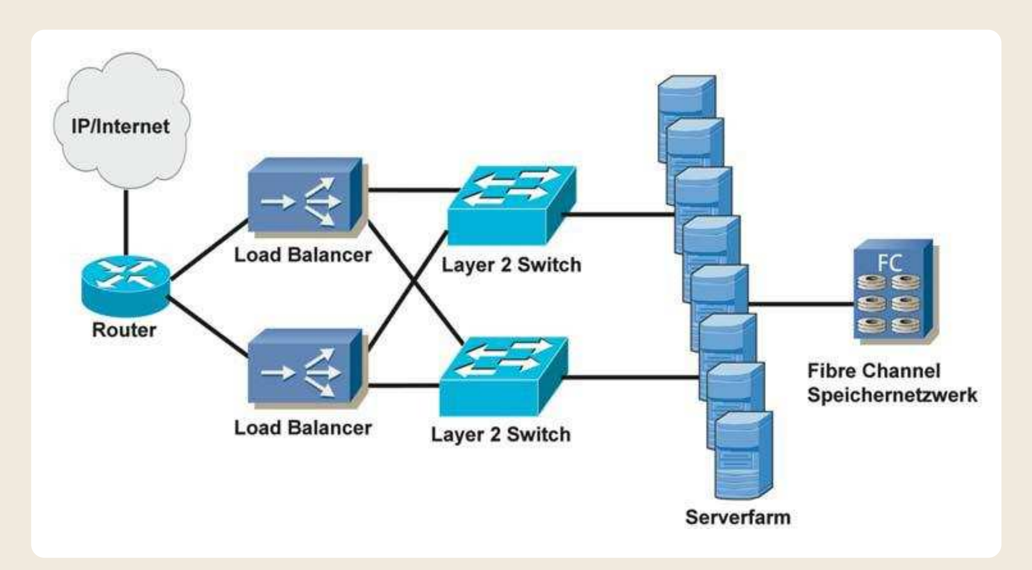
Will man ein oder mehrere Geräte vor die Serverfarm spannen, die die Anfragen entgegennehmen und sie gleichmäßig an die Server verteilen, greift man zu sogenannten Load Balancern. Sie haben ihren Ursprung in der Bereitstellung von Internetdiensten.

In der Regel legt man sie redundant aus, damit sie keinen „Single Point of Failure“ (SPoF) bilden. Sie können entweder als Hot-Standby konfiguriert sein – hier übernimmt das Zweitgerät nur bei einem Ausfall automatisch die Arbeit – oder im Parallelbetrieb. Dass beide Geräte gleichzeitig arbeiten, erlauben jedoch nicht alle Systeme. Einige Hersteller setzen dafür identische Geräte voraus.

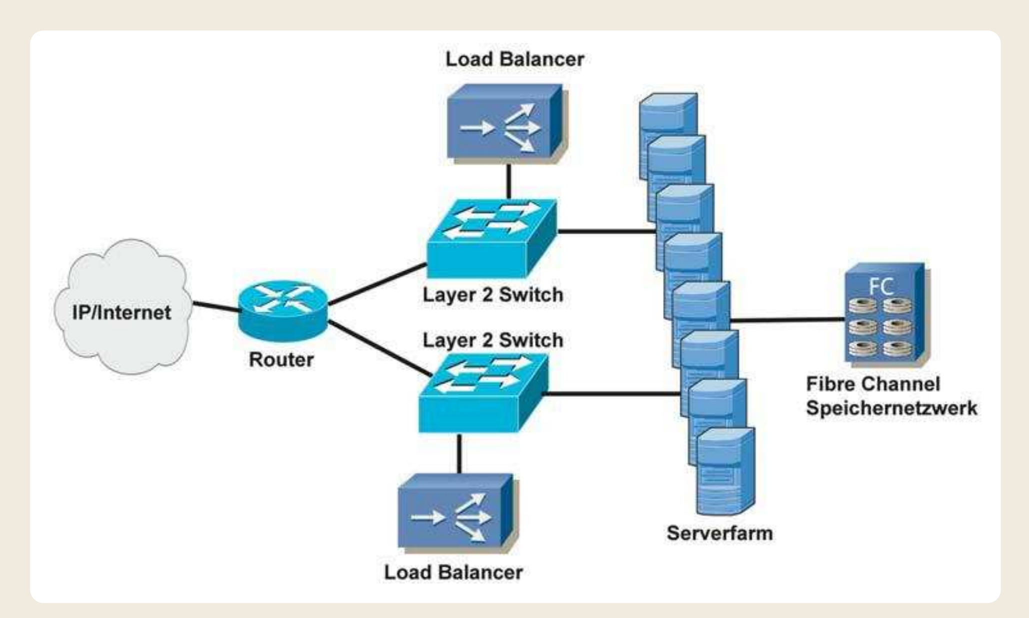
Im einfachen Fall verteilt der Load Balancer die Anfragen statisch, meist reihum (Round Robin). Insbesondere für Server mit identischer Hardwareausstattung ist dies eine akzeptable Methode. Als Alternative gibt es die Aufteilung nach einem vorgegebenen Schlüssel. Server unterschiedlicher Leistungsfähigkeit lassen sich durch eine einfache Gewichtung berücksichtigen, etwa im Verhältnis 2:1 (Weighted Round Robin).

Erzeugen alle Transaktionen etwa die gleiche Last auf dem Server, bietet sich das Least-Connection-Verfahren an. Es teilt die Anfrage jeweils dem Server zu, der gerade die wenigsten Verbindungen bedient.

Oft stellt jedoch nicht jede Anfrage die gleichen Anforderungen an Bearbeitungszeit und Ressourcen. Effektiver arbeiten dynamische Methoden: Eine bessere Performance erreicht man, wenn der Load Balancer



Wer für eine große Anzahl von Servern viele Ports benötigt, die der Load Balancer nicht direkt zur Verfügung stellen kann, schaltet weitere Layer-2-Switches davor (Abb. 2).



Sollen die Daten auf dem Rückweg den Load Balancer umgehen, müssen die Server direkt über einen Layer-2-Switch erreichbar sein. Beide Geräte lassen sich in einem Chassis vereinen (Abb. 3).

Kenntnis über die Auslastung der einzelnen Server hat und die nächste Anfrage jeweils dem Server zuteilt, der am wenigsten beschäftigt ist. Dazu müssen die Server den Load Balancer über ihren Auslastungsgrad in Kenntnis setzen. Im einfachsten Fall ermittelt ein *ping* denjenigen Server, der am schnellsten antwortet, weil er vermutlich über die meisten freien Ressourcen verfügt (Least Response Time). Gleichzeitig lässt sich damit die generelle Verfügbarkeit, also auch ein eventueller Ausfall des Servers, feststellen.

Weitaus bessere Ergebnisse erzielt man jedoch mit speziellen Programmen (Agents) auf den Servern, die Informationen sammeln und an den Load Balancer entweder über die bestehende Netzverbindung oder eine zusätzliche serielle Leitung senden: CPU- und Speicherauslastung, aber auch die Verfügbarkeit spezieller Ressourcen und Daten. Damit kann der Load Balancer die Anfragen gezielter verteilen. Einige Load Balancer verwenden dafür das Standard-Netzmanagement-Protokoll SNMP (Simple Network Management Protocol).

Verteilungsgerechtigkeit

Eine Methode, zu der Load Balancer gern greifen, ist die Network Address Translation (NAT), die (viele) interne IP-Adressen in eine oder mehrere externe umsetzt. Sie wurde entwickelt, um der Adressenknappheit der heute noch geläufigen IP-Version 4 entgegenzuwirken. So lassen sich mehr IP-Adressen verwenden, als der Internet-Provider zuteilt. NAT Load Balancing setzt die individuellen IP-Adressen mehrerer Server in eine einzige um, sodass nach außen nur ein logischer Server erscheint. Anders als bei NAT in Routern arbeiten Load Balancer aber nicht mit einer eindeutigen Zuordnung, sondern mit einem Algorithmus, der die Anfragen lastbezogen auf die Server verteilt. Der Vorteil liegt in der einfachen Implementierung, der Nachteil darin, dass der Load Balancer den Flaschenhals bildet, denn auch alle zurückgesandten Pakete passieren ihn zwecks IP-Adressumwandlung – vor allem, wenn der Rückweg für die Performance vieler Anwendungen wie Downloads oder Videoübertragung entscheidend ist.

Diesen Engpass umgeht ein Verfahren namens MAC Address Translation (MAT). Dabei lässt der Load Balancer die IP-Adressen unverändert und tauscht stattdessen seine MAC-Adresse (Media Access Control) gegen die des ausgewählten Servers aus. Leitet er das Paket weiter, besteht eine scheinbar direkte Kommunikation zwischen Client und Server, während der Server die Daten unter Umgehung des Load Balancer zurücksendet. Dieses Verfahren nennt sich Flat Based Load Balancing und bedingt, dass Server und Load Balancer über einen Switch verbunden sind, der eine direkte Verbindung zwischen Server und Client unter Umgehung des Load Balancer erlaubt (siehe Abbildung 3).

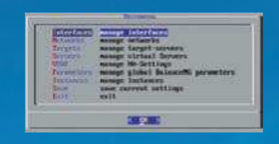
Ein Load Balancer setzt keineswegs voraus, dass auf allen angeschlossenen Servern die gleiche Anwendung läuft, also beispielsweise ein Webserver. Bestimmte Serveranwendungen wie Webdienste, E-Mail, FTP, Spiele oder Datenbanken können auf dedizierten Servern laufen, während ein Load Balancer die Anfragen nach der jeweiligen Anwendung verteilt. Dazu zieht er den Zielport des gewünschten Dienstes heran (Network Address Port Translation oder NAPT). In solchen Fällen firmieren Load Balancer gern unter dem Begriff „Content Switch“. Die einfache Aufgabe erfordert streng genommen

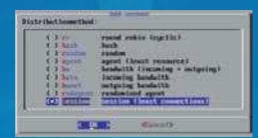


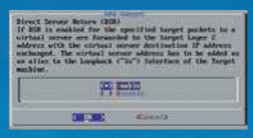
Klassische Load-Balancer-Appliances im 19-Zoll-Format wie die von Team ix kommen mit wenigen Ports aus (Abb. 4).

LOADBALANCER

KEINE ANGST VOR LOADBALANCING









BalanceNG® certified system

Mit dem Thomas-Krenn LoadBalancer können Sie schnell und einfach ein professionelles loadbalancing selbst in vorhandene Strukturen integrieren und damit sowohl Performance als auch Erreichbarkeit Ihrer Anwendungen erhöhen. **www.thomas-krenn.com/loadbalancer**

AB EUR 2.299,-

NEU





DER THOMAS-KRENN LOADBALANCER

- Aufeinander abgestimmte Hardware-/Software- Kombination
- Schnelle Einrichtung durch menübasiertes Konfigurationstool
- Master/Backup-Hochverfügbarkeit mit VRRP nach rfc3768 (min. 2 LoadBalancer notwendig)
- Protokollunabhängiges Load Balancing
- Unterstützung von „Direct Server Return“ (DSR)
- bis zu 128 Instanzen, 1024 Targets und 512 virtuelle Server möglich

www.thomas-krenn.com · EU: +49 (0) 8551 9150-0 · AT: +43 (0) 7282 20797-3600 · CH: +41 (0) 848 207970 **Made in Germany!**

nicht einmal einen Load Balancer: Ein Router mit NAPT und Port Forwarding reicht bereits aus. Der kann die Anfragen allerdings nur nach Diensten und nicht nach Last verteilen.

Auch die Aufteilung der Webpräsenz auf mehrere Server beherrschen einige Load Balancer. Sie erkennen dann anhand der Verzeichnispfade in der URL, welcher Server für die Anfrage zuständig ist. So lassen sich lokale und zentrale, statische und dynamische Inhalte oder auch einfach nur die schiere Größe eines Webauftrittes über mehrere Server streuen.

Neben der reinen Verteilung kann der Load Balancer eine Priorisierung vornehmen, also Datenpakete bevorzugt behandeln, wenn sie ein Echtzeitverhalten erfordern (etwa Sprache) oder zu einer wichtigen Anwendung gehören. Beispielsweise bietet es sich an, HTTPS gegenüber HTTP zu priorisieren, um Bestell- und Zahlungsvorgänge zügig abzuwickeln.

Hoch hinaus

Für viele Anwendungen reichen die genannten einfachen Algorithmen der Lastverteilung nicht aus – etwa wenn der Load Balancer logisch zusammengehörige Datenpakete identifizieren und dem richtigen Server zuordnen muss. Viele Webanwendungen wie E-Commerce oder Onlinebanking arbeiten transaktionsorientiert, was bedeutet, dass sie eine gewisse Anzahl elementarer Transaktionen zusammenhängend ausführen müssen, etwa Auswahl der Ware, Eingabe der Lieferadresse und Überprüfung der Kreditkarte. Dazu stellen sie eine logische Verbindung zwischen Client und Server her, die sogenannte Session oder Sitzung, die für den gesamten Vorgang gilt.

Der Load Balancer muss sicherstellen, dass eine Session komplett auf einem Server läuft. Dazu muss er tief in den Inhalt der Datenpakete vordringen (Deep Packet Inspection). Es gibt unterschiedliche Informationen, die der Load Balancer dafür heranziehen kann, Datenpakete einer offenen Session zuzuordnen, etwa die IP-Adresse des Absenders, eine Session- oder SSL-ID oder Cookies.

Mit Hash-Verfahren kann man das Speichern und Suchen der Quell-IP-Adressen umgehen. Über eine Hash-Tabelle ordnet der Load Balancer dabei die Client-IP-Adresse dem Zielserver zu. Entspricht die Quelladresse im Datenpaket nicht der wirklichen Adresse des Clients, weil ein Proxy oder NAT-Router vorgeschaltet ist, läuft die Maßnahme ins Leere.

Fällt ein Server während einer Session aus, hängt es von der Anwendung ab, ob ein anderer Server sie übernehmen kann. Der Load Balancer kann zwar die Datenpakete einfach an einen anderen Server weiterleiten, aber der Anwender merkt nur dann nichts davon, wenn Anwendung und Daten es erlauben. Andernfalls muss er die letzte Aktion wiederholen.

Neben den Session-Informationen muss der Load Balancer je nach Anwendung eine Reihe anderer Informationen auswerten, beispielsweise Cookies, Browser- und Gerätetypen oder Teile der URL. Hinzu kommt, dass E-Commerce- und Banking-Sessions in der Regel mit SSL (Secure Socket Layer) oder TLS (Transport Layer Security) verschlüsselt ablaufen, was zusätzliche Anforderungen an die Kapazität des Load Balancer stellt. Da er die relevanten Informationen aus verschlüsselten Datenpaketen nicht ermitteln kann, muss der Load Balancer als Endpunkt der Verschlüsselung fungieren. Einige Hersteller bieten hierfür zusätzliche Hardware zur Erhöhung der Leistungsfähigkeit an.



Viel Eisen für große Server: Foundrys Server Iron bietet eine hohe Portdichte und nutzt bei Bedarf 10-Gigabit-Ethernet (Abb. 5).

Zwei links, zwei rechts

Während Layer-2-Switches nur wenige Bytes (vornehmlich die Adressen) zum Weiterleiten des Pakets analysieren müssen, benötigen Load Balancer weitere Informationen, die aus höheren Protokollschichten stammen. Deshalb bezeichnet man sie auch als Layer-4-7- oder einfach als Layer-7-Switches. Sie müssen stets das gesamte Paket zwischenspeichern und auswerten und benötigen daher deutlich mehr Rechenkapazität als Switches oder Router (siehe auch iX extra 1/2006 [6]).

Für eine gute Performance setzen viele Hersteller auf schnelle Switch-Hardware mit entsprechenden Puffern für die Datenpakete, die Basisfunktionen besonders schnell ausführen kann. Darauf installieren sie Spezialsoftware für die Zusatzfunktionen. Obwohl sie Geschwindigkeitsvorteile bieten, finden sich heute kaum noch Load Balancer am Markt, die nur auf Layer 4 arbeiten. Grund sind die immer schnellere Hardware und der Wunsch nach mehr Funktionsumfang und Flexibilität.

Da der Load Balancer für seine Arbeit bereits große Teile aus dem Header und dem Datenteil der Pakete analysieren muss, fällt es leicht, an dieser Stelle Erweiterungen anzusiedeln, etwa Sicherheitsfunktionen wie Access Control, Intrusion Detection, eine Firewall oder Traffic Management – und das auch auf Anwendungsebene. Derlei Funktionen haben Load Balancern auch den Namen „Reverse Proxy“ eingetragen, weil sie ähnlich wie ein Proxy-Server agieren, um Performance und Sicherheit zu erhöhen – nur eben nicht vor den Clients, sondern vor den Servern.

Heute bedeutet Load Balancing Skalierung und Ausfallsicherheit für zahlreiche Serveranwendungen, also auch für Voice over IP (etwa in IP-TK-Anlagen) oder Video-Streaming. Es bietet darüber hinaus die direkte Unterstützung von Datenbanken oder großer Anwendungen wie SAP oder Oracle. Aber nicht nur Anwendungen im eigentlichen Sinn nutzen Load Balancing, sondern auch eine Reihe von Diensten aus der Internet-Infrastruktur wie DNS, DHCP und Radius, die Terminierung von VPN-Tunneln oder Lastverteilung zwischen Firewalls.

Neben dem großen Anwendungsgebiet Internet spielt das Load Balancing auch bei firmeninternen Client-Server-Anwendungen eine Rolle, einem Thema, das mit zunehmendem Einsatz von Thin Clients eine Renaissance erlebt.

Leistungsparameter für Load Balancer sind die Anzahl der gleichzeitig handhabbaren Sessions (gegebenenfalls mit SSL-Verschlüsselung), die Bearbeitungszeit pro Datenpaket sowie



Quelle: Cisco

Cisco integriert Lastverteilungs-Funktionen in Service-Module für die Catalyst-Switches und nennt sie ACE wie Application Control Engine (Abb. 6).

Anzahl und Durchsatz der verfügbaren Ports. Darüber hinaus unterscheiden sie sich hinsichtlich ihrer Arbeitsweise auf Layer 4 bis 7, der Unterstützung spezieller Anwendungen sowie möglicher Zusatzfunktionen, zum Beispiel Security.

Hard- oder Software?

Load Balancer können recht unterschiedlich ausgeführt sein und im einfachen Fall auf Standardhardware und -software aufsetzen, wie etwa die Appliance von Zeus, die aus Linux-gesteuerten PC-Komponenten besteht. Das ermöglicht eine kostengünstige Lösung und hohe Flexibilität. Erweiterungen wie Sicherheitsfunktionen lassen sich einfach verwirklichen. Auch reine Softwarelösungen sind erhältlich, beispielsweise Virtual Appliance von Loadbalancer.org oder BalanceNG von Inlab. Eine Reihe von Herstellern nutzt als Softwaregrundlage Linux Virtual Server, eine frei verfügbare Software für Linux-Server. Mit ausreichender Kenntnis kann man sie auf einer Serverfarm installieren, die unter Linux läuft (www.linuxvirtualserver.org). Die offenen Schnittstellen erlauben eine einfache Anpassung und Erweiterung, etwa durch den Einsatz kundenspezifischer Skripte.

Eine höhere Performance lässt sich mit Spezialhard- und -software erzielen, bis hin zu dafür optimierten ASICs. Entsprechend den immer komplexeren Anwendungen und zusätzlichen Anforderungen finden sich kaum noch Geräte, die sich auf reines Load Balancing beschränken, und auch die Hersteller rücken dies nicht mehr in den Vordergrund. F5 etwa rüstet seine Viprion-Serie mit einer Reihe zusätzlicher Funktionen aus und verkauft sie als Application Delivery Controller (ADC); Cisco wiederum implementiert die Funktionen auf einem Board in den modularen Layer-2-Switches und nennt es Application Control Engine (ACE). Für kleinere Unternehmen ist diese Funktion aber auch in einem separaten Chassis erhältlich.

Seit der Übernahme der Technikpioniere Alteon durch Nortel und Arrowpoint durch Cisco gehört Load Balancing zu den Produktportfolios der großen Hersteller von Netzwerkgeräten. Dort läuft es aber meist unter Schlagworten wie Content Delivery Network (CDN) oder Application Acceleration. Das zeigt, dass Load Balancing eine ergänzende Technik zur Beschleunigung und Skalierung von Serveranwendungen ist, die sich je nach Anwendung durch andere ergänzen oder ersetzen lässt. Einige Hersteller, etwa Radware, bilden die Funktionen in separaten Geräten (Appliances) ab, andere als separate Module oder in eine große Box integriert.

Weitere Bestandteile von CDNs sind Caching oder Streaming-Verfahren wie Multicast. Erstere dienen vor allem der Beschleunigung von Downloads, indem Kopien der Daten in lokalen oder global verteilten Caches nah am Client liegen und dazu beitragen, dass entsprechende Anforderungen oft gar nicht bis zum eigentlichen Server gelangen müssen. Beim Multicasting fürs Audio- und Video-Streaming baut nicht jeder Client eine exklusive Verbindung zum Server auf. Vielmehr nutzen mehrere Clients Teile der Datenleitungen gemeinsam. Andere CDN-Techniken sind zum Beispiel Traffic Flow Management und Datenkompression. Je nach konkretem Angebot verstehen die Hersteller unter CDN auch nur einzelne der genannten Aspekte, etwa das Caching. Manchmal bezieht sich der Begriff „Load Balancing“ nur auf Layer 4, in Unterscheidung zu CDN auf Layer 7.

Mit zunehmender Auslagerung von Servern aus dem eigenen Rechenzentrum zu Hostern und der Entwicklung zu Produkten wie Managed Server (siehe iX extra 9/2007 [7]) nimmt auch die Zahl der Load Balancer in den Rechenzentren der Hoster zu. Neben der einfachen Form des Unterstellens der Geräte als Co-Location- oder Rackspace-Angebote mit eigener Verantwortung bieten Hoster zunehmend Lösungen zum Management der gesamten Hardware einschließlich Load Balancer, Firewall et cetera an. Das entlastet den Auftraggeber von den technischen Details, erfordert aber eine genaue Spezifikation der Leistungsparameter, die es zu erfüllen gilt. (hw/sun)

Uwe Schulze *ist Fachautor in Berlin.*

Literatur

[1] Marc Grimme, Mark Hlawatschek; Dateisysteme; Aus dem Pool schöpfen; Global File System Storage Cluster 6; *iX* 11/2004, S. 118

[2] Hansjörg Maurer; Clusterdateisysteme; Simultanz; GFS und GPFS unter Linux im Vergleich; *iX* 10/2005, S. 86

[3] Udo Seidel; Dateisysteme; Gruppentherapie; Oracles Cluster File System; *iX* 8/2007, S. 118

[4] Johann Baumeister; Netzwerktechnik; Unter dem Joch; Load Balancing: Add-ons, Appliances, Software; *iX* 8/2005, S. 114

[5] Uwe Schulze; Netzwerke; Domain, sweet Home; Eigene Internet-Adressen registrieren; *iX* extra 5/2007, S. I

[6] Uwe Schulze; Netzwerke; Schalter mit Zukunft; Vom Hub zum Switch: Die Entwicklung des Ethernet; *iX* extra 1/2006, S. V

[7] Uwe Schulze, Daniel Wetzel; Netzwerke; Boom-Branche Hosting; Vom Webspace zum eigenen Server; *iX* extra 9/2007, S. VIII

Balanceakt

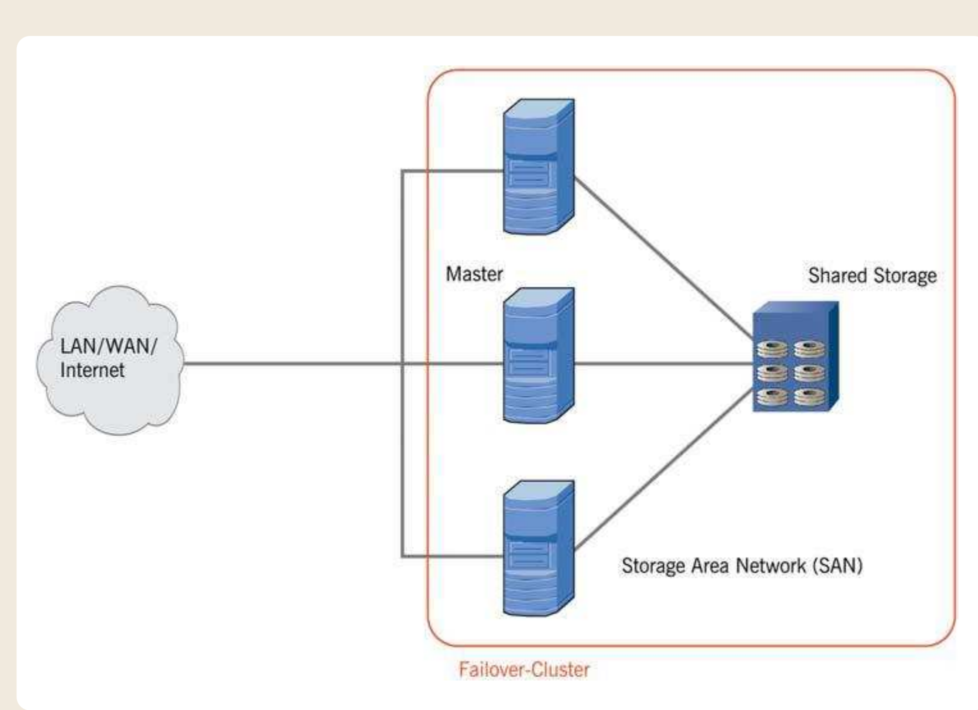
Software für Load-Sharing- und Failover-Cluster

Eine Alternative zur Load-Balancer-Appliance sind Softwarepakete zur Lastverteilung. Will man das Nadelöhr eines vorgeschalteten Servers oder Geräts vermeiden, kann man die Server auch direkt zusammenschalten – zum Load-Sharing- oder Failover-Cluster.

Wer nur zwei oder drei Server zusammenschalten will, wird sich erst einmal – schon aus Kostengründen – nach einer geeigneten Software für das Load Balancing umschauen. Dabei findet man solche, mit der sich Standardhardware in einen vorgeschalteten Load Balancer verwandeln lässt, und solche, die direkt auf den beteiligten Servern läuft und dort die Aufgaben verteilt. Außerdem gibt es eine Reihe von Betriebssystemen und Distributionen wie OpenBSD und Red Hat Cluster Server, die die Lastverteilung von sich aus beherrschen. Microsoft hat seinen Cluster Service (MSCS) und das Network Load Balancing (NLB) ebenfalls in einige Versionen der Windows Server 2003 und 2008 integriert.

Beim integrierten Konzept – ohne vorgeschalteten Load Balancer – sprechen die Hersteller gern von „Cluster-Lösungen“, was schnell zu einer Verwechselung mit dem sogenannten HA-Cluster führt (siehe Artikel „Des anderen Last“, S. I). Zumindest eins haben beide gemeinsam: Jeder Cluster erscheint nach außen als ein einziger logischer Rechner; seine Mitglieder (Knoten) haben eine „gemeinsame Identität“ (siehe Abbildung 1).

Zudem benutzt ein Active/active-Rechnerverbund denselben Überwachungsmechanismus wie ein HA-Cluster (active/passive). Sogenannte Watchdogs halten die anderen Mitglieder des Clusters per Heartbeat über den eigenen Gesundheitszustand und gegebenenfalls über den Auslastungsgrad auf dem Laufenden. Verstummen die Lebenszeichen, läuft nach wenigen Sekunden eine Konvergenz genannte Routine ab, die den ausgefallenen Rechner aus dem Cluster entfernt respektive vom Stromnetz trennt. In den meisten Fällen übernimmt ein Knoten die interne Verwaltung des Clusters: der Master-Knoten. Fällt der aus, stimmen in größeren Rechner-Verbänden die verbleibenden Mitglieder per „Quorum“ darüber ab, wer der neue Master-Knoten wird.



Wie jeder Cluster erscheint auch der Failover-Cluster, dessen Knoten hier auf ein gemeinsames Speichersystem zugreifen, als ein logischer Rechner im Netz (Abb. 1).

Einfach verteilen

Anders als bei HA-Clustern tritt der Verfügbarkeitsaspekt bei Load-Sharing-Clustern in den Hintergrund. Beim Ausfall eines Knotens verteilen sich neue Anfragen auf die verbliebenen Mitglieder. Dieser Mechanismus ist jedoch nicht für jede Anwendung geeignet, vor allem wenn Sessions unterbrochen werden, Informationen verloren gehen und die Benutzer ihre Programme erneut starten müssen. Das Hauptanwendungsgebiet der Load-Sharing-Cluster sind Webserver – wie bei den vorgeschalteten Load Balancern. Darüber hinaus kommen derartige Cluster beispielsweise als Streaming-Media-, Terminal-, VPN- und Proxy-Server sowie als Firewall vor.

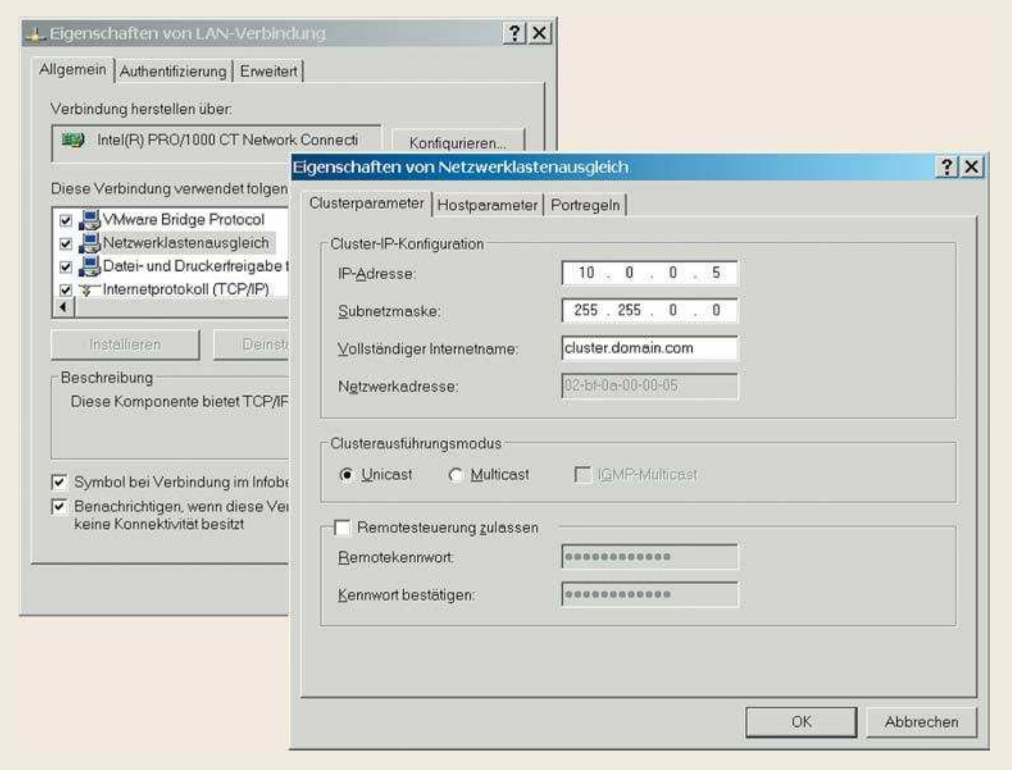
Ein Beispiel dafür ist das Network Load Balancing von Microsoft. Es kennt zwei Betriebsarten: Unicast und Multicast (siehe Abbildung 2). Im Unicast-Modus erhalten alle Hosts im Cluster dieselbe virtuelle MAC-Adresse (Media Access Control). Auf diese Weise können sie die für den Cluster bestimmten Datenpakete parallel in Empfang nehmen; anschließend entscheidet jeder Host selbst, ob ein Paket für ihn bestimmt ist oder nicht. Damit sowohl der Austausch von Heartbeats als auch die Konvergenz problemlos funktionieren, kommunizieren die Hosts über ein separates Netz, für das sie eine zweite Netzwerkkarte besitzen.

Dagegen behalten die Hosts im Multicast-Modus ihre eigene MAC-Adresse und kommen dadurch mit einer Netzwerkkarte aus. Multicasting hat jedoch den Nachteil, dass eine Reihe von Routern und Switches damit nicht zurechtkommt und sie deshalb ein Nachkonfigurieren von Hand erfordern. Der Admin nimmt im Router einen

SOFTWARE ZUR LASTVERTEILUNG UND AUSFALLSICHERHEIT

Die folgende Übersicht erhebt keinen Anspruch auf Vollständigkeit.

Bezeichnung	Betriebssystem	Anwendungsgebiet	URL
2X Load Balancer for Terminal Services/Citrix	Windows	vorgeschaltetes Load Balancing für Terminalserver	www.2x.com/loadbalancer
Apache Module mod_proxy_balancer	alle Posix-Plattformen (BSD, Linux, Mac OS X, Unix), Windows	vorgeschaltetes Load Balancing für Apache-Webserver	httpd.apache.org/docs/2.2/mod/mod_proxy_balancer.html
Beowulf	Linux	HPC-Cluster	www.beowulf.org
Citrix Netscaler Global Load Balancing for Presentation Server	Windows	XenApp-Farmen	www.citrix.com/English/ps2/products/subfeature.asp?contentID=21770
Crossroads	alle Posix-Plattformen (BSD, Linux, Mac OS X, Unix)	vorgeschaltetes Load Balancing	crossroads.e-tunity.com
DRBD	Linux	HA-Cluster	www.drbd.org
Fastream Technologies IQ Reverse Proxy	Windows	vorgeschaltetes Load Balancing für Webserver	www.iqreverseproxy.com/en/iqreverseproxy.php
Gnu Queue (Pre-Alpha)	Linux	vorgeschaltetes Load Balancing	www.gnu.org/software/gnu-queue
HAproxy	Linux, Solaris, Posix-Plattformen	vorgeschaltetes Load Balancing für Webserver	haproxy.1wt.eu
Heartbeat	Linux	HA-Cluster	www.linux-ha.org
HP Serviceguard	Linux	Failover- und Load-Sharing-Cluster	www.hp.com/solutions/enterprise/highavailability/linux/serviceguard
IBM HACMP for System p (High Availability Cluster Multi-Processing)	Linux (Power), AIX	HA-Cluster	www.ibm.com/systems/p/software/hacmp/index.html
IBM Parallel Sysplex	z/OS	HA- bis Load-Sharing-Cluster	www.ibm.com/systems/z/advantages/pso/index.html
Inlab BananceNG	Linux, Solaris	vorgeschaltetes Load Balancing	www.inlab.de/balanceng
Keepalived	Linux	HA für Linux Virtual Server	www.keepalived.org
Linux Virtual Server	Linux	Failover- und Load-Sharing-Cluster	www.linux-vs.org
Loadbalancer.org Virtual Appliance	Windows, Mac, Linux	vorgeschaltetes Load Balancing	loadbalancer.org/download.html
Microsoft Cluster Service	Windows Server	Failover-Cluster	support.microsoft.com/kb/q259267
Microsoft Network Load Balancing	Windows Server	Load-Sharing-Cluster	support.microsoft.com/kb/240997
Novell Cluster Services	Netware 5	Failover- und Load-Sharing-Cluster	www.novell.com/products/clusters/ncs101/details.html
OpenBSD	OpenBSD	Load-Sharing-Cluster	www.openbsd.org/41.html
OpenBSD relayd	OpenBSD	vorgeschaltetes Load Balancing	www.openbsd.org/cgi-bin/cvsweb/src/usr.sbin/relayd
Oracle Real Application Cluster (RAC)	HP-UX, Oracle-Linux, Solaris, Windows	HA- bis Load-Sharing-Cluster	www.oracle.com/database/rac_home.html
Pound – Reverse-Proxy and Load Balancer	Linux, OpenBSD, Solaris, andere Posix-Plattformen (nicht getestet)	vorgeschaltetes Load Balancing für Webserver	www.apsis.ch/pound/index_html
Pramati Web Load Balancer	Pramati Server	vorgeschaltetes Load Balancing für Webserver	www.pramati.com/index.jsp?id=pro_psv350012
Prestwood Load Balancer	Windows	vorgeschaltetes Load Balancing für Webserver	www.prestwood.com/products/lb
Red Hat Cluster Suite	Red Hat Enterprise Linux	HA-Cluster	www.redhat.com/cluster_suite
Resonate Central Dispatch	AIX, HP-UX, Linux, Solaris, Windows Server	vorgeschaltetes Load Balancing	www.resonate.com/prod_centr_disp.html
XLB HTTP Load Balancer	alle Posix-Plattformen (BSD, Linux, Mac OS X, Unix)	vorgeschaltetes Load Balancing für Webserver	sourceforge.net/projects/xlb
Zeus ZXTM Software	Linux, Solaris, FreeBSD	vorgeschaltetes Load Balancing	www.zeus.com/products/zxtm_software



Microsofts Network Load Balancing hier auf einem Windows Server 2003 – kann im Unicast- oder im Multicast-Modus arbeiten (Abb. 2).

statischen ARP-Eintrag (Address Resolution Protocol) vor, der die gemeinsame Cluster-IP-Adresse mehreren MAC-Adressen zuordnet. Um diesen Aufwand zu sparen, arbeitet Microsofts Network Load Balancing standardmäßig im Unicast-Modus.

Stets zu Diensten

Eine höhere Ausfallsicherheit als reine Load-Sharing-Cluster bieten solche mit Failover-Funktion. Sie arbeiten mit den im vorangehenden Beitrag beschriebenen Cluster-Dateisystemen (siehe Abbildung 1) auf externen RAID-Subsystemen, meist über ein SAN angebunden. Typische Vertreter für Failover-Cluster sind HPs Serviceguard für Linux, Novells Cluster Services für Netware 5 und Microsofts Failover Cluster Service für Windows Server 2008. Dank gemeinsamer Nutzung des Shared Storage sind die Cluster-Knoten in der Lage, die Aufgaben eines ausgefallenen Knotens unmittelbar zu übernehmen und auf die zuletzt genutzten Daten zuzugreifen. Im Idealfall bemerken die Nutzer davon nichts oder höchstens eine um einige Sekunden verzögerte Reaktion des Systems. Vor allem können sie sich darauf verlassen, dass keine Daten verloren gehen. Failover-Cluster stellen in Datenbanken, Messaging-Systemen und Terminaldiensten sowie in Datei- und Druckdiensten die Skalierbarkeit und eine hohe Verfügbarkeit sicher.

Allerdings erfordern Failover- einen höheren Aufwand als reine Load-Sharing-Cluster. Außer dem gemeinsam genutzten Speicher besitzen sie entweder zusätzliche Knoten, die bei Störungen in Aktion treten, oder in der Kalkulation ihrer I/O- und Rechenkapazität ist der Ausfall eines Knotens mit berücksichtigt (N+1). Die Knoten, die noch miteinander kommunizieren können, müssen dazu aushandeln, wer von ihnen weiterhin im Cluster verbleiben darf und wer die Master-Funktion übernimmt. Dazu schreiben alle verbliebenen Knoten nach einem ausgefeilten Schema auf eine „Voting Disk“. Und sie müssen feststellen, ob andere ehemalige Cluster-Mitglieder versuchen, auf ihr Dateisystem zuzugreifen. Für dieses Prozedere sind in der Regel mehrere Neustarts der beteiligten Rechner nötig. (hw/sun)

Dagmar Hemke
ist IT-Beraterin in Berlin.